

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
11. Januar 2001 (11.01.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/02590 A2

- (51) Internationale Patentklassifikation⁷: C12N 15/82, (74) Anwälte: BETTENHAUSEN, Berthold usw.; Dehmel & Bettenhausen, Müllerstr. 1, D-80469 München (DE).
15/29, C07K 14/415, A01H 5/00
- (21) Internationales Aktenzeichen: PCT/DE00/02233 (81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (22) Internationales Anmeldedatum:
3. Juli 2000 (03.07.2000)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
199 30 570.6 2. Juli 1999 (02.07.1999) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. [DE/DE]; Hofgartenstrasse 8, D-80539 München (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): SAGASSER, Martin [DE/DE]; Lichtstrasse 23, D-50825 Köln (DE). WEIS-SHAAR, Bernd [DE/DE]; Fingerhutweg 13, D-50226 Frechen (DE). DEKKER, Koen [NL/DE]; Goldammerweg 9, D-50829 Köln (DE).
- (84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: PLANTS WITH MODIFIED GENE EXPRESSION

(54) Bezeichnung: PFLANZEN MIT VERÄNDERTER GENEXPRESSION

(57) Abstract: The invention relates to a method for producing a plant with modified gene expression, comprising the stable integration of a seed-specific regulatory sequence or a fragment or derivative thereof and a nucleic acid sequence that is functionally linked to said seed-specific regulatory sequence or fragment or derivative and that codes for a gene product in the genome of plant cells or plant tissues; and the regeneration of the resulting plant cells or plant tissues to produce plants. The invention also relates to a method for producing plants with a modified flavonoid content, comprising the stable integration of at least one nucleic acid sequence according to SEQ ID NO:2 or 4 or a nucleic acid sequence that is homologous with this, or a fragment or derivative thereof in the genome of plant cells or plant tissues, and the regeneration of the resulting plant cells or plant tissues to produce plants.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung einer Pflanze mit veränderter Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen. Die vorliegende Erfindung betrifft ferner ein Verfahren zur Herstellung von Pflanzen mit verändertem Flavonoidgehalt, umfassend das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat in das Genom von Pflanzenzellen oder Pflanzengewebe und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

WO 01/02590 A2

Pflanzen mit veränderter Genexpression

Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung einer Pflanze mit veränderter
10 Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen
Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen
Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt
kodierenden Nukleinsäuresequenz in das Genom von Pflanzenzellen oder Pflanzengewebe und
15 Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen. Die vorliegende
Erfindung betrifft ferner ein Verfahren zur Herstellung von Pflanzen mit verändertem
Flavonoidgehalt, umfassend das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß
SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment
oder Derivat in das Genom von Pflanzenzellen oder Pflanzengewebe und Regeneration der
erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

20

Der pflanzliche Phenylpropanoidstoffwechsel, zusammenfassend beschrieben z.B. von
Weisshaar & Jenkins, Current Opinion in Plant Biology 1, 251-257 (1998) oder von Shirley,
Trends in Plant Sciences 1, 377-382 (1996), besteht aus einem Netz sich verzweigender
biochemischer Reaktionsketten, die die Pflanze mit den verschiedensten phenolischen
25 Komponenten versorgen. Der identische Beginn aller nachfolgenden Synthesen geht von
Phenylalanin aus und führt über Cinnamat und 4-Coumarat zu Coumaroyl-CoA. Beteiligt sind
die Enzyme Phenylalanin-Ammonium-Lyase (PAL), Cinnamat-4-Hydroxylase (C4H) und 4-
Coumarat-CoA-Ligase (4CL). Im weiteren Verlauf verzweigen sich die Synthesen. Eine
wichtige Verzweigung führt zur Biosynthese von Flavonoiden. Flavonoide sind Flavanderivate,
30 d.h. ausschließlich bei Pflanzen vorkommende Sekundärmetaboliten mit dem zwei aromatische
Ringe beinhaltenden Flavan-Grundgerüst. Zu den vielen, oft spezifisch in einer bestimmten
Pflanzenart vorkommenden Flavonoiden zählen unter anderem UV-absorbierende Flavonole, zu
den Gerbstoffen gehörende Tannine und z.B. als rote und blaue Blütenfarbstoffe gebildete
Anthocyane.

35

In *Arabidopsis thaliana* sind verschiedene chromosomale Gen-Loci bekannt, die eine Rolle in der Flavonoidbiosynthese spielen. Mutationen in etlichen dieser Loci verhindern die Akkumulation brauner Farbstoffe in der Samenschale (Testa) und werden als transparent testa (tt in mutierter Form, TT als Wildtyp) bezeichnet. Durch die unter der Samenschale liegenden
5 Kotyledonen erscheint der Samen in diesen Fällen gelblich bis hellbraun. Wildtypsamens sind dagegen dunkelbraun. Einige dieser Loci (tt3, tt4, tt5, ttg) sind außerdem noch an der Produktion von Anthocyanen in Blättern und Sproß beteiligt und ein Locus (ttg) hat eine zusätzliche Funktion bei der Entwicklung von Trichomen und Wurzelhaaren. Alle bisher aufgetretenen tt-Mutanten haben sich als rezessiv erwiesen. Eine zusammenfassende Beschreibung findet sich in
10 Shirley et al., The Plant Journal 8, 659-671 (1995).

Die ersten Arabidopsis-Mutanten mit Defekt in der Flavonoidbiosynthese wurden 1971 von Bürger beschrieben (Bürger, Arabidopsis Information Service 8, 36-42 (1971)). In dieser Publikation wurde der Phänotyp von tt1, der eine abweichende Samenfarbe aufweist, erstmals
15 erwähnt. Durch genetische und morphologische Untersuchungen von Koornneef (Koornneef, Arabidopsis Information Service 18, 45-51 (1981), Koornneef, Arabidopsis Information Service 27, 1-4 (1990) sowie Shirley (Shirley et al., The Plant Journal 8, 659-671 (1995) konnte der Genlocus von tt1 auf Chromosom 1 – 54,9 sowie die Beschränkung des Phänotyps auf den Samen festgestellt werden.

20

Bei vielen Nutz- und Zierpflanzen ist die Veränderung der Samenzusammensetzung und eine Veränderung der Flavonoidbiosynthese aus landwirtschaftlichem und produktionstechnischen Gesichtspunkten seit langem erwünscht. Eine begrenzte Einflußnahme auf Komponenten des Flavonoidstoffwechsels war, solange die Genstruktur der beteiligten Enzyme nicht bekannt war,
25 nur mit den Mitteln der klassischen Züchtung möglich. Diese konventionelle Methode, beruhend auf der zufälligen Vermischung des väterlichen und mütterlichen Erbguts, ist relativ zeit- und kostenintensiv. Das gewünschte Ergebnis ist erst nach 10 bis 15 Jahren zu erwarten. Die gezielte Manipulation einzelner Komponenten der Flavonoidbiosynthese ist mit klassischer Züchtung ohne genetische Analysen kaum zu erreichen. Es besteht somit die Aufgabe, Pflanzen mit einer
30 Veränderung der Zusammensetzung des pflanzlichen Samens und einer Verbesserung der Samenqualität von Pflanzen z.B. Nutz- und Zierpflanzen bereitzustellen.

Die Aufgabe wird durch den Gegenstand der Patentansprüche gelöst.

Die Erfindung wird durch die folgenden Figuren erläutert.

Figur 1 zeigt eine schematische Darstellung der Analyse der Nukleinsäuresequenz des TT1-Promoters durch Vergleich mit TRANSFAC MATRIX TABLE, Rel.3.3 (E. Wingender et al., 1998). Dabei sind Bindestellen für Transkriptionsfaktoren in Form von Großbuchstaben (SBF-1 like sites: siehe Lawton et al., Silencer region of a chalcone synthase promoter contains multiple binding sites for a factor, SBF-1, closely related to GT-1, Plant Molecular Biology 16, 235-249 (1991)), kursive Schrift (AGAMOUS like sites: Huang et al., Isolation and characterization of the binding sequences for the product of the Arabidopsis floral homeotic gene AGAMOUS, Nucleic Acids Research 21, 4769-4776 (1993)), unterstrichenen Buchstaben (P like sites: Grotewold et al., The myb-homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset, Cell 76, 543-553 (1994)), großen und kursiven Buchstaben (MYB Ph3 like sites: Solano et al., Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.PH3) from Petunia hybrida, EMBO Journal 14, 1773-1784 (1995)) sowie großen und unterstrichenen Buchstaben (Athab-1 und 2 like sites: Sessa et al., The athb-1 and-2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities, EMBO Journal 12, 3507-3517 (1993)) gekennzeichnet. Das Start-ATG ist fett und unterstrichen dargestellt. Die Numerierung beginnt mit der 5' gelegenen SpeI Schnittstelle im verwendeten Plasmidvektor pSK-TT1.

20

Figur 2 zeigt die Nukleinsäuresequenz der genomischen DNA-Sequenz von TT1, beginnend mit dem Start-ATG. Großbuchstaben stellen dabei Exons und kursiv geschriebene Buchstaben Introns dar. Die Numerierung setzt die Numerierung aus Figur 1 fort.

25 Figur 3 zeigt die für das TT1-Gen von Arabidopsis kodierende cDNA-Sequenz und die daraus abgeleitete Aminosäuresequenz von TT1.

Figur 4 zeigt schematisch einen Vergleich der TT1-Aminosäuresequenz mit Sequenzen aus der NCBI GenBank. Acc.No AL049660, AB025629 und AC006085.9 sind aus Nukleinsäuresequenzen abgeleitete hypothetische Aminosäuresequenzen aus *Arabidopsis thaliana* (At), AJ234704 ist eine aus einer Nukleinsäuresequenz abgeleitete hypothetische Aminosäuresequenz für *Hordeum vulgare* (Hv). In der Konsensussequenz bezeichnet ein ! Aminosäuren vom Typ I oder V, ein \$ Aminosäuren vom Typ L oder M, ein % Aminosäuren

30

vom Typ F oder Y sowie ein # Aminosäuren vom Typ N, D, Q, E, B oder Z.

Figur 5 zeigt eine schematische Darstellung des Restriktionsmusters von pSK-TT1.

- 5 Figur 6 zeigt eine Darstellung der Samenfärbung der Mutante *ttl* im Vergleich zum Wildtyp.

Figur 7 zeigt eine Darstellung der nukleäre Lokalisation von TT1.

- Figur 8 zeigt X-Gluc gefärbte Blüten und Schoten transgener *TT1*-GUS-Pflanzen in
10 verschiedenen Entwicklungsstadien. A zeigt eine Blüte kurz nach der Befruchtung. Die GUS-Aktivität ist auf die apikalen Samenanlagen beschränkt. Ältere Blüten (B) und junge Schoten (F) zeigen GUS-Aktivität in Funiculi und Integumenten apikaler und distaler Samenanlagen (C und D). Mit der verwendeten Methode ist GUS-Aktivität in Samenanlagen, die sich nicht mehr weiter entwickeln, nicht aber in älteren Samenstadien nachweisbar (E).

15

- Der hier verwendete Ausdruck „homologe Sequenz“ oder „homologe Nukleinsäuresequenz“ oder "Homolog" bezeichnet eine Nukleinsäure- oder Proteinsequenz mit signifikanter Ähnlichkeit zur Vergleichssequenz oder Teilen davon, wobei diese homologen Sequenzen eine Aktivität oder Teilaktivität vergleichbar der Aktivität der Nukleinsäure- oder Proteinsequenzen
20 gemäß SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 oder SEQ ID NO:4 aufweisen. Als homologe Sequenzen gelten Nukleinsäuresequenzen, die mit den Sequenzen gemäß SEQ ID NO:1, SEQ ID NO:2 oder SEQ ID NO:4 oder Teilen dieser Sequenzen unter stringenten oder wenig stringenten Bedingungen hybridisieren (zu stringenten und wenig stringenten Bedingungen siehe Sambrook et al., Molecular Cloning, Cold Spring Harbour Laboratory
25 (1989), ISBN 0-87969-309-6). Ein Beispiel für stringente Hybridisierungsbedingungen ist: Hybridisierung in 4 x SSC bei 65° C (alternativ in 50% Formamid und 4 X SSC bei 42° C), gefolgt von mehreren Waschschritten in 0,1 x SSC bei 65°C für insgesamt eine Stunde. Ein Beispiel für wenig stringente Hybridisierungsbedingungen ist Hybridisierung in 4 x SSC bei 37° C, gefolgt von mehreren Waschschritten in 1 x SSC bei Raumtemperatur. Als homologe Sequenzen
30 sollen des weiteren Nukleinsäure- oder Proteinsequenzen oder Teile davon gelten, die unter Zuhilfenahme des Similaritätsalgorithmus BLAST (Basic Local Alignment Search Tool, Altschul et al., Journal of Molecular Biology 215, 403-410 (1990) eine signifikante Ähnlichkeit mit den Nukleinsäure- und Aminosäuresequenzen der vorliegenden Erfindung aufweisen. Als

signifikant ähnlich werden, wie hier verwendet, Sequenzen bezeichnet, die z.B. unter Verwendung von Standardparametern im Blast-Service des NCBI ein Signifikanzniveau (Probability) von $P < 10^{-5}$ aufweisen, wenn Sie mit den Sequenzen gemäß SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 oder SEQ ID NO:4 oder Teilen davon verglichen werden.

5

Der hier verwendete Ausdruck "Derivate" bezeichnet Nukleinsäuresequenzen, die eine oder mehrere Deletionen, Substitutionen, Additionen, Insertionen und/oder Inversionen aufweisen.

Der hier verwendete Ausdruck "funktional verbunden" bedeutet, daß eine regulatorische Sequenz wie ein Promotor die Expression eines Gens steuert oder daß eine Nukleinsäuresequenz von dem Promotor ausgehend exprimiert wird.

Der hier verwendete Ausdruck "Vektor" bezeichnet natürlich vorkommende oder künstlich erschaffene Konstrukte zur Aufnahme, Vermehrung, Expression oder Übertragung von Nukleinsäuren, z.B. Plasmide, Phagemide, Cosmide, künstliche Chromosomen, Bakteriophagen, Viren, Retroviren.

Der hier verwendete Ausdruck "Expressionssystem" bezeichnet jedwede Kombination von Vektoren, Restriktionsenzymen, Transformationsmethoden, Zellextrakten, lebenden Zellen z.B. prokaryotischen oder eukaryotischen Zellen oder Organismen mit dem Zweck der endogenen oder exogenen Expression von Genen.

Die vorliegende Erfindung betrifft die regulatorische samenspezifische Nukleinsäuresequenz (im folgenden auch Promotor bezeichnet), die natürlicherweise in *Arabidopsis thaliana* die Expression des TT1-Gens steuert. Die erfindungsgemäße Nukleinsäuresequenz ist in SEQ ID NO:1 aufgeführt. Ferner betrifft die Erfindung ein Fragment oder Derivat der Nukleinsäuresequenz gemäß SEQ ID NO:1 oder eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert und für die samenspezifische Expression verantwortlich ist. Bevorzugt ist eine Nukleinsäuresequenz, die unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert.

30

Die erfindungsgemäße Nukleinsäuresequenz kann natürlichen Ursprungs sein oder künstlich hergestellt worden sein. Die Nukleinsäuresequenz kann sowohl in Sense- als auch in Antisense-

Orientierung vorliegen.

Die erfindungsgemäße Nukleinsäuresequenz eignet sich z.B. zur Identifizierung und Isolierung von zum TT1-Gen homologer Gene in anderen Organismen oder von homologen Genen in
5 *Arabidopsis thaliana* mit Hilfe spezieller Hybridisierungs- oder Screening-Verfahren, z.B. als Sonde für das Screening in DNA-Bibliotheken mit Hilfe der Hybridisierung an einzelsträngige Nukleinsäuren ähnlicher Basenabfolge.

Die erfindungsgemäße Nukleinsäuresequenz eignet sich ferner zur spezifischen Steuerung der
10 Expression von Genen in Organismen oder Zellen, bevorzugt zur spezifischen Steuerung der Expression von Genen in Samen, insbesondere in der Samenschale. Bevorzugt betrifft die vorliegende Erfindung ein Verfahren zur Herstellung einer Pflanze mit veränderter Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen
15 Sequenz oder deren Fragment oder Derivat funktional verbundenen mit einer für ein Genprodukt codierenden Nukleinsäuresequenz in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen. Die mit der erfindungsgemäßen Nukleinsäuresequenz funktional verbundenen Nukleinsäuren können endogene, exogene genomische DNA-Abschnitte oder cDNAs oder deren Fragmente oder
20 Derivate sein. Endogen bedeutet dabei, daß die Nukleinsäuresequenz aus dem gleichen Organismus stammt, in den sie mit dem erfindungsgemäßen Verfahren integriert wird, z.B. eine Nukleinsäuresequenz aus *Arabidopsis thaliana* wird mit dem erfindungsgemäßen Verfahren in *Arabidopsis thaliana* integriert. Exogen bedeutet, das die Nukleinsäuresequenz aus einem anderen Organismus stammt, z.B. eine Nukleinsäuresequenz aus *Arabidopsis thaliana* wird mit
25 dem erfindungsgemäßen Verfahren in z.B. Weizen integriert. Die Nukleinsäuresequenzen können gegenüber den natürlich vorkommenden Nukleinsäuresequenzen Deletionen, Substitutionen, Additionen, Insertionen und/oder Inversionen aufweisen.

Die erfindungsgemäße Nukleinsäuresequenz kann in Vektoren, Expressionssystemen oder
30 Pflanzen, Pflanzengewebe oder Pflanzenzellen oder tierischen Zellen oder Mikroorganismen zur Veränderung der Expressionsmuster verschiedenster Genprodukte verwendet werden. Die Expression der Genprodukte kann dabei gegenüber ihrer natürlichen Expression sowohl verstärkt als auch verringert sein.

Beispielsweise kann die erfindungsgemäße Nukleinsäuresequenz für die Expression des samenschalenspezifischen TT1-Gens verwendet werden. Ferner eignet sich die erfindungsgemäße Nukleinsäuresequenz auch für die Regulation der Expression anderer Gensequenzen für jedwede Anwendung, sowohl aus *Arabidopsis thaliana* als auch aus anderen
5 Organismen. Dabei kann der Promotor in Kombination mit beliebigen Genen vorliegen, sowohl in einem Vektor als auch in transgenen Organismen.

Insbesondere kann die erfindungsgemäße Nukleinsäuresequenz zur Kontrolle der Expression weiterer natürlicher samenspezifischer oder künstlich in den Samen transferierter Gene
10 eingesetzt werden, z.B. zur Expressionsregulation weiterer Gene des Phenylpropanoidstoffwechsels, z.B. Phenylalanin-Ammonium-Lyase, Cinnamat-4-Hydroxylase, 4-Coumarat-CoA-Ligase, Chalcon-Synthase, Chalcon-Isomerase, Chalcon-Reduktase, Flavanon 3-Hydroxylase, Flavonoid-3'-Hydroxylase, Flavonoid-3'5'-Hydroxylase, Dihydroflavonol-4-Reduktase, Leucoanthocyanidin-Reduktase, Leucoanthocyanidin-Dioxygenase, 3'-
15 Glucosyltransferase, 5'-Glucosyltransferase, O-Methyl-Transferase. Die Kontrolle der Expression von Genen des Phenylpropanoidstoffwechsels mit Hilfe der erfindungsgemäßen Nukleinsäuresequenz eignet sich dabei insbesondere z.B. zur Verstärkung der UV-Absorptionsrate, Veränderung der Farbe, Verbesserung des Geschmacks oder der Lagerfähigkeit, Verstärkung des Schutzes vor Schädlingen oder Verbesserung der
20 Verarbeitungsfähigkeit verschiedener Pflanzengewebe.

Ferner kann die erfindungsgemäße Nukleinsäuresequenz zur Expressionsregulation von Genen, kodierend für weitere Proteine der Samenschale, verwendet werden. Die Proteine können die Qualität, d.h. die Zusammensetzung der Samen verbessern und/oder deren physiologische
25 Eigenschaften verändern. Beispiele für bevorzugte Proteine der Samenschale sind Proteine, die die Keimstimmung oder Dormanz beeinflussen. Die Keimruhe von Samen (Dormanz) wird in vielen Fällen durch verschiedene Inhaltsstoffe der Samenschale verursacht bzw. gesteuert. So kann eine Veränderung der Dormanz erreicht werden, indem die erfindungsgemäße Nukleinsäuresequenz zur Regulation von in der Samenschale exprimierten und die
30 Keimstimmung oder Dormanz beeinflussenden Genen eingesetzt wird. Solche Gene sind z.B.

- a) Gene, die an der Bildung von wasserundurchlässigen Schichten (z.B. wachshaltige Cuticulae oder Suberinlamellen) beteiligt sind, z.B. Genbank Acc. No. AF030260 (Cytochrom P450 CYP94A1), Genbank Acc. No. M80567 (Lipid transfer protein LTP1),

- b) Gene für die Synthese von Samenschalenkomponenten, die dem Embryo einen mechanischen Widerstand entgegensetzen wie z.B. Lignine, z.B. Genbank Acc. No. J02979 (Lignin forming peroxidase),
- c) Gene für Proteine, die den mechanischen Widerstand der Samenschale gegen den Embryo schwächen, wie z.B. Zellwandbestandteile abbauende Enzyme, z.B. Genbank Acc. No. 5 AJ242807 (Cellulase), Genbank Acc. No. AJ277900 (beta 1,3-Glucanase),
- d) Gene für die Synthese von Wachstumsinhibitoren, wie z.B. Abscisinsäure, z.B. Genbank Acc. No. U95953 (viviparous 14), Genbank Acc. No. AF190462 9-(cis-epoxycarotenoid dioxygenase),
- 10 e) Gene für die Synthese von Samenschalenkomponenten, die Wachstumsinhibitoren wie z.B. Abscisinsäure in der Samenschale zurückhalten,
- f) Gene für die Synthese von Samenschalenkomponenten, die den Gasaustausch und damit die Sauerstoffversorgung des Embryos beeinflussen,
- g) Gene für Komponenten des Sekundärstoffwechsels, die die Vitalität des Samens 15 beeinflussen.

Weitere Beispiele für bevorzugte Proteine der Samenschale sind Proteine, die direkt oder indirekt Resistenz der Samen gegen Pathogenbefall durch Insekten, Pilze, Bakterien, Viren oder Nematoden vermitteln. An entsprechenden pflanzlichen Abwehrmechanismen sind verschiedene 20 Klassen von Proteinen und Sekundärmetaboliten beteiligt. In diesem Sinne kann eine verbesserte Pathogenresistenz des Samens erreicht werden, indem die erfindungsgemäße Nukleinsäuresequenz zur Regulation von in der Samenschale exprimierten und an der Pathogenabwehr beteiligten Genen eingesetzt wird. Solche Gene sind z.B.

- a) Gene für insektizid wirksame α -Amylase-Inhibitoren, Proteinase-Inhibitoren und 25 Faserproteine, z.B. Genbank Acc. No. D26109 (alpha-amylase inhibitor-2), Genbank Acc. No. AF105340 (proteinase inhibitor precursor),
- b) Gene für die Synthese polymerer Zellwandbestandteile wie z.B. Kallose, die als physische Barriere gegenüber Pilz- und Bakterieninfektionen dienen, z.B. Genbank Acc. No. AF085717 (callose synthase catalytic subunit)
- 30 c) Gene für hydrolytische Enzyme wie z.B. Glucanasen und Chitinasen, die die Zellwände des Pathogens auflösen, z.B. Genbank Acc. No. AF241267 (chitinase 2),
- d) Gene für die Synthese antimikrobiell wirksamer Phytoalexine, z.B. Genbank Acc. No. U69554 (6a-hydroxymaackiain methyltransferase)

- e) Gene aus der Gruppe der pflanzlichen R-Gene (Resistenz-Gene), deren Genprodukte direkt oder indirekt über weitere Proteine mit Genprodukten von avr-Genen (Avirulenz-Genen) des Pathogens wechselwirken und zum programmierten Zelltod infizierter Pflanzenzellen an der Infektionsstelle („Hypersensitive Response“) führen, z.B. Genbank Acc. No. BE039015 (downy mildew resistance protein gene rpp5), Genbank Acc. No. AF122994 (RPM1 variant gene), AF098962, Genbank Acc. No. AF122994 (RPM1 variant gene), Genbank Acc. No. AF098962 (disease resistance protein RPP1-WsA gene)
- 5 f) Gene, kodierend für Proteine mit DNA-bindenden WRKY-Domänen, die durch Bindung an sogenannte W-Boxen Pathogenabwehr-Gene aktivieren, z.B. Genbank Acc. No. AF193770 (WRKY 3), Genbank Acc. No. AF193771 (WRKY 4),
- 10 g) Gene für Saponine, die Pilzmembranen durch Bindung an Sterole zerstören können.

Des weiteren ist die erfindungsgemäße Nukleinsäuresequenz zur Verbesserung des Nährwertes und der Verdaulichkeit des Samens geeignet. Eine entsprechende Veränderung kann z.B. durch

15 eine Verringerung des Rohfasergehaltes, eine Reduzierung an antinutritiven Substanzen oder eine Veränderung des Gehaltes an Proteinen und Speicherlipiden erreicht werden, indem die erfindungsgemäße Nukleinsäuresequenz zur Regulation von an diesen Prozessen beteiligten Genen eingesetzt wird.

- 20 Außerdem ist die erfindungsgemäße Nukleinsäuresequenz zur Regulation der Speicherung von Reservestoffen im Samen, z.B. von Stärke, geeignet. Möglichkeiten der Einflußnahme auf Reservestoffe bestehen z.B. durch die Expression von Sense- oder Antisense-Transkripten des Kohlenhydratstoffwechsels der Pflanze wie z.B. ADP-Glucose-Synthetase, Stärkesynthase, ADP-Glucose-Pyrophosphorylase oder die Expression von Genen anderer Organismen wie z.B.
- 25 Hefe-Invertase zur Mobilisierung der Stärke.

- Die erfindungsgemäße Nukleinsäuresequenz eignet sich ferner zur Verlagerung von Stoffwechselprodukten des Samens, z. B. von Glucosinolaten zur Verbesserung der Nematodenresistenz, in die Samenschale. Ferner kann der Promotor zur Verhinderung oder
- 30 Verzögerung der Reifung der Samenschale durch kontrollierte Expression von Ribonuklease-Genen benutzt werden. Eine Verzögerung der Reifung der Samenschale kann zur Bildung von größeren Samen führen. Des weiteren kann die Entfernung der Samenschale durch eine verzögerte Reifung erleichtert werden.

Geeignete Vektoren zur Aufnahme und Überführung der erfindungsgemäßen Nukleinsäuresequenz und der mit ihr funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz können die Vermehrung und/oder die Expression der aufgenommenen Nukleinsäuren in Einzellern wie z.B. *Escherichia coli* oder *Agrobacterium tumefaciens* oder in
5 Pflanzenzellen, Pflanzengewebe oder Pflanzen oder tierischen Zellen oder Tieren gewährleisten. Entsprechende Vektoren können natürlich vorkommen oder aber künstlich hergestellt sein. Die Vektoren können Selektionsmarker, Terminatorsequenzen, Polylinker, Promotorelemente, Enhancer, Polyadenylierungsstellen und andere genetische Elemente umfassen. Zur Klonierung geeignete Vektoren sind z.B. pBluescript, Plasmide der pUC-Serie,
10 Plasmide der pGem-Reihe oder auf dem Bakteriophagen λ basierende Vektoren. Ein zur Verwendung in *Agrobacterium* benutzter Plasmidvektor ist z.B. pBin19 (Bevan et al., *Nucleic Acids Research* 12, 8711-8721. (1984)). Zur Transformation und Expression in Pflanzen stellen auf dem Ti-Plasmid von *Agrobacterium*-Arten oder auf Pflanzenviren aufbauende Konstrukte verwendungsfähige Vektoren dar und sind dem Fachmann bekannt. Eine zusammenfassende
15 Beschreibung bislang benutzter Vektoren findet sich in Guerinéau und Mullineaux, *Plant Transformation and Expression Vectors*, in: *Plant Molecular Biology Labfax*, herausgegeben von Croy, Oxford, BIOS Scientific Publishers, 121-148 (1993).

Einige der in handelsüblichen Transformations- und Expressionssystemen angewendeten
20 Transformationsmethoden zur Übertragung von Fremdgenen (Transformation) in das Genom von Pflanzen werden im folgenden vorgestellt. Die Wahl der Methode zur Einbringung der erfindungsgemäßen Nukleinsäuresequenz und der mit ihr funktional verbundenen für ein Genprodukt kodierenden Nukleinsäuresequenzen in pflanzliche Zellen ist jedoch nicht auf diese Liste beschränkt. Bislang eingesetzte Transformationsverfahren bei Pflanzen sind z.B. der
25 Gentransfer mittels *Agrobacterium tumefaciens* (z.B. durch Baden von Samen oder Blattstückchen in einer Agrobakterienlösung), mittels pflanzlicher Viren, durch Elektroporation, durch Einschießen (microprojectile bombardment) oder Einspritzen (Mikroinjektion) sowie die Inkubation von trockenen Embryonen in DNA-haltigen Flüssigkeiten und die Transformation von Protoplasten unter Zuhilfenahme von Polyethylenglykol. Genauere Beschreibungen der
30 angesprochenen Verfahren finden sich z.B. in Jens et al., *Techniques for Gene Transfer*, in: *Transgenic Plants, Vol. 1, Engineering and Utilization*, herausgegeben von Kung und Wu, Academic Press 128-143 (1993).

Die erfindungsgemäße Nukleinsäuresequenz kann zur Kontrolle der Genexpression in Mikroorganismen wie z.B. *Escherichia coli* oder *Saccharomyces cerevisiae*, in mono- und dikotylen Pflanzen sowie Algen und tierischen Zellen und Tieren verwendet werden. Besonders bevorzugte Pflanzen sind dabei Kulturpflanzen wie z.B., Soja, Reis, Baumwolle, Zuckerrübe, 5 Raps, Rübsen, Sonnenblume, Flachs, Hanf, Kartoffel, Tabak, Tomate, Luzerne, Salat, Erbse, Bohne, Karotte, Zwiebel, die verschiedenen Baum, Nuß- und Weinspezies, ferner Getreide wie z.B. Gerste, Weizen, Roggen, Hafer, Triticale, Mais, Futtergräser, ferner Obstsorten wie z.B. Mango, Apfel, Pfirsich, Stachelbeere, Johannisbeere, Banane, Melone, Kürbis und diverse Zitrusfrüchte wie z.B. Zitrone, Apfelsine, Pampelmuse, Mandarine. Die mit der 10 erfindungsgemäßen Nukleinsäuresequenz transformierten Pflanzen können unmodifizierte Wildtyp-Pflanzen oder durch Züchtung erhaltene Pflanzen oder modifizierte Pflanzen z.B. transgene Pflanzen sein. Die erfindungsgemäße Nukleinsäuresequenz kann ferner nicht nur in Pflanzen oder Pflanzengeweben, sondern auch in einzelnen Pflanzenzellen, z.B. in einer Zellkultur, verwendet werden.

15

Die vorliegende Erfindung betrifft ferner eine transformierte Zelle, insbesondere eine transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, in der die erfindungsgemäße Nukleinsäuresequenz und die damit funktional verbundene für ein Genprodukt kodierenden Nukleinsäuresequenz stabil integriert sind. Ferner betrifft die 20 vorliegende Erfindung eine mit der erfindungsgemäßen Nukleinsäuresequenz transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, die oder das zu einer samenproduzierenden Pflanze regenerierbar ist. Insbesondere betrifft die vorliegende Erfindung eine Pflanze, die nach dem erfindungsgemäßen Verfahren erhältlich ist. Ferner betrifft die vorliegende Erfindung Saatgut, das von Pflanzen erhalten wird, die nach dem 25 erfindungsgemäßen Verfahren erhalten werden.

Die vorliegende Erfindung betrifft ferner transgene Pflanzen mit einer stabil in das Genom integrierten samenspezifischen regulatorischen Nukleinsäuresequenz gemäß SEQ ID NO:1, oder deren Fragment oder Derivat oder Homolog mit der biologischen Funktion eines 30 samenspezifischen Promotors, und einer mit dieser Nukleinsäuresequenz funktionell verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz entsprechend den vorstehend aufgeführten und beschriebenen Beispielen für solche Gene.

Die vorliegende Erfindung betrifft ebenfalls die genomische Sequenz oder die cDNA-Sequenz des TT1-Gens von *Arabidopsis thaliana* Ecotyp Columbia, das über die Bildung von Zwischenprodukten für die Bildung von Flavonoiden verantwortlich ist. Die erfindungsgemäße Nukleinsäuresequenz ist in SEQ ID NO:2 und 4 aufgeführt. Ferner betrifft die Erfindung ein

5 Fragment oder Derivat der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert und für die Bildung von Flavonoiden verantwortlich ist. Bevorzugt ist eine Nukleinsäuresequenz, die unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert. Die vorliegende Erfindung betrifft ferner die Aminosäuresequenz des TT1-Gens

10 von *Arabidopsis thaliana* Ecotyp Columbia gemäß SEQ ID NO:3.

Die erfindungsgemäße Nukleinsäuresequenz eignet sich z.B. zur Identifizierung und Isolierung zum TT1-Gen homologer Gene in anderen Organismen oder von homologen Genen in *Arabidopsis thaliana* mit Hilfe spezieller Hybridisierungs- oder Screening-Verfahren, z.B. als

15 Sonde für das Screening in DNA-Bibliotheken mit Hilfe der Hybridisierung an einzelsträngige Nukleinsäuren ähnlicher Basenabfolge.

Die vorliegende Erfindung betrifft ferner ein Verfahren zur Herstellung von Organismen oder Zellen, insbesondere von Pflanzen, mit verändertem Flavonoidgehalt, umfassend das stabile

20 Integrieren mindestens der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat in das Genom von Zellen, insbesondere von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

25 Die erfindungsgemäße Nukleinsäuresequenz oder deren Fragment oder Derivat kann natürlichen Ursprungs sein oder aber künstlich hergestellt worden sein. Die Nukleinsäuresequenz oder deren Fragment oder Derivat kann sowohl in Sense- als auch in Antisense-Orientierung verwendet werden. Ferner kann die erfindungsgemäße Nukleinsäuresequenz oder deren Fragment oder Derivat in den genomischen Locus des TT1-Gens oder eines zum TT1-Gen homologen Gens von

30 *Arabidopsis thaliana* oder in einen genomischen Locus eines zum TT1-Gen homologen Gens einer anderen Pflanze durch homologe Rekombination integriert sein. Des weiteren kann die Nukleinsäuresequenz auch in Form von Ribonukleinsäuren z.B. als Ribozym verwendet werden. In diesem Fall sind die Thymin-Basen (T) durch Uracil-Basen (U) ersetzt. Die Bildung von

Flavonoiden kann durch die erfindungsgemäße Nukleinsäuresequenz sowohl verstärkt als auch verringert sein.

Die erfindungsgemäße Nukleinsäuresequenz kann z.B. zur Expression in Mikroorganismen z.B. Escherichia coli oder Saccharomyces cerevisiae und in mono- und dikotylen Pflanzen sowie Algen und tierischen Zellen und Tieren verwendet werden. Bevorzugte Pflanzen sind dabei Kulturpflanzen z.B., Soja, Reis, Baumwolle, Zuckerrübe, Raps, Rüben, Sonnenblume, Flachs, Hanf, Kartoffel, Tabak, Tomate, Luzerne, Salat, Erbse, Bohne, Karotte, Zwiebel, die verschiedenen Baum, Nuß- und Weinspezies, ferner Getreide wie z.B. Gerste, Weizen, Roggen, Hafer, Triticale, Mais, Futtergräser, sowie Obstsorten z.B. Mango, Apfel, Pfirsich, Stachelbeere, Johannisbeere, Banane, Melone, Kürbis und diverse Zitrusfrüchte z.B. Zitrone, Apfelsine, Pampelmuse, Mandarine. Die mit der erfindungsgemäßen Nukleinsäuresequenz transformierten Pflanzen können unmodifizierte Wildtyp-Pflanzen oder durch Züchtung erhaltene Pflanzen oder modifizierte Pflanzen z.B. transgene Pflanzen sein. Die erfindungsgemäße Nukleinsäuresequenz kann ferner nicht nur in Pflanzen oder Pflanzengeweben sondern auch in einzelnen Pflanzenzellen z.B. in einer Zellkultur verwendet werden.

Eine Expression der erfindungsgemäßen Nukleinsäuresequenz kann durch Kombination der Sequenz mit einem geeigneten Promotor erreicht werden. Der in dieser Kombination vorliegende Promotor kann dabei sowohl der TT1-Promotor gemäß SEQ ID NO:1 als auch ein anderer endogener Promotor der transformierten Zelle oder ein auf dem Vektor befindlicher exogener Promotor sein. Als Promotor ist dabei grundsätzlich jede regulatorische Sequenz geeignet, die die Expression von Fremdgenen in Zellen, insbesondere in Pflanzen steuern kann, z.B. der CaMV 35S-Promotor aus dem Blumenkohl-Mosaik-Virus (Franck et al., Cell 21, 285-294 (1980)). Eine Expression der erfindungsgemäßen Nukleinsäuresequenz kann auch durch einen chemisch induzierbaren Promotor erreicht werden. Beispiele für chemisch induzierbare Promotoren sind der PRPI-Promotor (Ward et al., Plant Molecular Biology 22, 361-366 (1993)), ein durch Salizylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzensulfonamid induzierbarer Promotor (EP-A 388186), ein durch Tetrazyklin induzierbarer Promotor (Gatz et al., Plant Journal 2, 397-404 (1992)), ein durch Abscisinsäure induzierbarer Promotor (EP-A 335528) sowie ein durch Ethanol oder Cyclohexanon induzierbarer Promotor (WO 93/21334). Je nach gewünschtem Expressionsort können auch Promotoren verwendet werden, die in bestimmten Pflanzengeweben oder Pflanzenteilen aktiv sind. Beispiele für entsprechende

Promotoren sind der Phaseolin-Promotor (US 5504200), der Isoflavon-Reduktase Promotor (US 5750399), ein samenspezifischer Promotor aus Tabak (US 5824863) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al., EMBO Journal 8, 2445-2452 (1989)).

- 5 Ferner ist die erfindungsgemäße Nukleinsäuresequenz kombinierbar mit Sequenzen, die ein Targeting in bestimmte Kompartimente der Pflanze sicherstellen, z.B. für Transitpeptide oder Teile davon kodierende Sequenzen. Des weiteren sind Sequenzen, kodierend für enzymatisch aktive oder antigen wirksame Proteine z.B. His-tag mit der obengenannten erfindungsgemäßen Nukleinsäuresequenz kombinierbar.

10

Ferner kann die erfindungsgemäße Nukleinsäuresequenz in Vektoren, Expressionssystemen oder Pflanzen, Pflanzengeweben oder Pflanzenzellen oder tierischen Zellen oder Mikroorganismen zur Veränderung des Expressionsmusters verwendet werden.

- 15 Die erfindungsgemäße Nukleinsäuresequenz eignet sich ferner in Kombination mit verschiedenen Promotoren zur Manipulation der phänotypischen und genotypischen Eigenschaften verschiedener Pflanzen oder Pflanzengewebe, z.B. zur Veränderung der Samenfarbe, z.B. zur ästhetischen Verbesserung verschiedener Pflanzenarten. Zierpflanzen mit z.B. ausgeschaltetem oder mutiertem TT1-Gen können optisch attraktive Varietäten darstellen.

20

- Die erfindungsgemäße Nukleinsäuresequenz kann ebenso zur Verstärkung der UV-Schutzfunktion der Samenschale durch Veränderung des Flavonoidgehaltes verwendet werden. Flavonoide zeigen Absorptionsmaxima im ultravioletten Bereich. Absorbierend wirken dabei die delokalisierten π -Elektronen der phenolischen Gruppen. Eine Erhöhung der
- 25 Flavonoidkonzentration im Samen kann eine drastische Verringerung der UV-induzierten Gewebeschäden bewirken. Auf diese Weise läßt sich z.B. die Keimungsrate des Saatgutes, vor allem in Regionen mit intensiver Sonnenbestrahlung, verbessern.

- Ferner kann die erfindungsgemäße Nukleinsäuresequenz zur Verstärkung der Schutzfunktion der
- 30 Samenschale gegen Schädlingsbefall durch Veränderung des Flavonoidgehaltes verwendet werden. Proanthocyanidine und andere phenolische Komponenten wirken fungizid, u.a. aufgrund enzyminhibitorischer Eigenschaften (Jambunathan et al., Journal of Agricultural and Food Chemistry 34, 425-429 (1986)). Eine Konzentrierung dieser Stoffe in der Samenschale kann die

Pathogenresistenz erhöhen.

Ferner kann die erfindungsgemäße Nukleinsäuresequenz zur Verbesserung der Geschmacksqualität des Samens und anderer Pflanzenteile infolge einer Veränderung des Flavonoidgehaltes, insbesondere des Gehaltes an kondensierten Tanninen, verwendet werden. Kondensierte Tannine haben einen großen Einfluß auf den Geschmack vieler Obst- und Gemüsesorten z.B. von Apfel, Kiwi oder Banane. Auf pflanzlichen Extrakten basierende Getränke z.B. Kaffee, Tee, Wein und Fruchtsäften werden ebenfalls in ihrer Geschmacksqualität von Tanninen geprägt.

10

Außerdem können durch den Einsatz der erfindungsgemäßen Nukleinsäuresequenz Verdaulichkeit und Nährwert sowie die Qualität der Proteinfraction des Samens verbessert werden. Dabei kann der veränderte Flavonoidstoffwechsel die Menge und Zusammensetzung an kontaminierenden sekundären Inhaltsstoffen beeinflussen. Infolge des reduzierten Anteils an phenolischen Komponenten und antinutritiven Substanzen sind sowohl reinere als auch optisch attraktivere hellere Proteinfractionen aus dem Samen zu gewinnen und der Rohfasergehalt kann reduziert werden.

Des weiteren kann die erfindungsgemäße Nukleinsäuresequenz aufgrund der auffälligen Gelbsamigkeit entsprechender transgener Pflanzen als optisch erkennbare genetische Markierung in der Pflanzenzucht eingesetzt werden. Insbesondere eignet sie sich als züchtungsbegleitendes Markergen zur Kennzeichnung von Pflanzen und Saatgut mit weiteren, phänotypisch nicht sichtbaren Modifikationen zur erleichterten Unterscheidbarkeit von z.B. transgenen und nicht-transgenen Genotypen.

25

Eine weitere Verwendung der erfindungsgemäßen Nukleinsäuresequenz stellt die Verbesserung der Verarbeitungsfähigkeit des Samens in industriellen Produktionsprozessen dar. Beispielsweise verursachen kondensierte Tannine in der Testa des Samens der Gerste unerwünschte Präzipitate während des Bierbrauprozesses (Shirley, Seed Science Research 8, 415-422 (1998)). Durch reduzierte Tanningehalte kann die Präzipitatbildung verhindert werden.

30

Die vorliegende Erfindung betrifft ferner eine transformierte Zelle, insbesondere eine transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, in der die

- erfindungsgemäße Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat stabil integriert ist. Ferner betrifft die vorliegende Erfindung eine mit der erfindungsgemäßen Nukleinsäuresequenz transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, die oder das zu einer samenproduzierenden Pflanze regenerierbar ist. Insbesondere betrifft die vorliegende Erfindung eine Pflanze, die nach dem erfindungsgemäßen Verfahren erhältlich ist. Ferner betrifft die vorliegende Erfindung Saatgut, das von Pflanzen erhalten wird, die nach dem erfindungsgemäßen Verfahren erhalten werden.
- 10 Die vorliegende Erfindung betrifft ferner transgene Pflanzen mit einer stabil in das Genom integrierten Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat mit der biologischen Aktivität eines Polypeptids codiert durch die Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4, und gegebenenfalls einer mit dieser Nukleinsäuresequenz funktional verbundenen regulatorischen
- 15 Nukleinsäuresequenz entsprechend den vorstehend aufgeführten und beschriebenen Beispielen für solche Promotoren.

BEISPIELE

- 20 Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt.

Beispiel 1: Allgemeine Klonierungsverfahren

- Die im Rahmen der vorliegenden Erfindung durchgeführten Klonierungsschritte wie z.B.
- 25 Restriktionsspaltungen, Agarose-Gelelektrophorese, Reinigung von Nukleinsäurefragmenten, Transfer von Nukleinsäuren auf Filtermaterialien, Transformation und Anzucht von Bakterienzellen usw. wurden wie bei Sambrook et al., Molecular Cloning, Cold Spring Harbour Laboratory (1989), ISBN 0-87969-309-6, durchgeführt.

30 Beispiel 2: Herstellung einer Knockout-Population von *Arabidopsis thaliana*

- Die vorliegende Erfindung wurde durch das Screening einer mit dem Transposon En-1/Spm (Pereira et al, EMBO Journal 5, 835-841 (1986)) mutagenisierten Knockout-Population von Pflanzen der Spezies *Arabidopsis thaliana* Ecotyp Columbia erhalten. Die Integration der

transposablen Elemente in Gene der Mutterpflanze führt häufig zum Ausfall der entsprechenden Genfunktion und in vielen Fällen zu einem vom Wildtyp unterscheidbaren Erscheinungsbild der betroffenen Pflanze. Zum Aufbau der Knockout-Population wurde das autonome En-1 Element aus *Zea mays* mittels *Agrobacterium tumefaciens* in *Arabidopsis* übertragen. Das entsprechende

5 Transposon tagging System ist in Cardon et al., *Plant Molecular Biology* 23, 157-178 (1993) beschrieben. Das verwendete Ti-Plasmid, pGV3850HPT::pkEn2, beinhaltet das komplette En-1 Element als Integrat. Zur Selektion von Hygromycin-resistenten Transformanten trägt dieser Vektor das HPT-Gen unter der Kontrolle des viralen CaMV 35S Promotors. Samen einer Transformante mit einer einzelnen T-DNA-Insertion wurden auf Hygromycin-haltigem Medium

10 ausgesät. Samen der so entstandenen, gegen Hygromycin-resistenten Pflanzen (T₂-Generation) wurden anschließend auf Kanamycin-haltigem Medium ausgesät. Bei den auf diese Weise selektierten Pflanzen (T₃-Generation) war das En1-Element aus der T-DNA heraus transponiert. Mittels PCR wurde diejenigen Pflanzen der T₄-Generation identifiziert, die ein oder mehrere transponierte En-1 Elemente, jedoch keine En-1 Elemente mehr in der T-DNA trugen.

15 Diese Pflanzen beinhalteten jedoch noch die T-DNAs ohne integrierte En-1 Elemente. Zur Erzeugung von En-1 positiven, T-DNA negativen Pflanzen wurde die T₄-Generation mit dem Wildtyp *Arabidopsis thaliana* Ecotyp Columbia gekreuzt und En-1 positive, T-DNA negative Pflanzen (S₀-Generation) durch PCR identifiziert. Samen jeder dieser Pflanzen wurden über 6 - 12 Generationen (bis zur S₆- bzw. S₁₂-Generation) hinweg vermehrt, bis insgesamt 3 000 Linien

20 mit insgesamt 15 000 unabhängigen En-1 Insertionen zur Verfügung standen (Wisman et al., *Plant Molecular Biology* 37, 989-999 (1998), Baumann et al., *Theoretical and Applied Genetics* 97, 729-734 (1998)).

Beispiel 3: Screening nach TT-Mutanten

25 Zur Identifizierung von phänotypisch auffälligen Mutanten der entstandenen En-1 Population wurden 2000 Familien der S₆- Generation (mit jeweils 20 Individuen) per Auge nach Auffälligkeiten durchmustert. Dabei konnte eine Linie (5K69) identifiziert werden, die sich durch eine abweichende Farbe der Samenschale (gelb statt dunkelbraun) auszeichnete, während die Produktion von Anthocyanen in Sproß und Blättern augenscheinlich nicht beeinträchtigt

30 war.

Die Beschränkung des Phänotyps auf den Samen wurde von Koornneef, supra, und Shirley, supra, auch für die klassische tt1-Mutante beschrieben. Um zu überprüfen, ob es sich bei der in

der En-Population gefundenen Linie um ein Allel von tt1 handelte, wurden beide Linien miteinander gekreuzt. Die Nachkommen dieser Kreuzung produzierten wiederum gelbe Samen. Das deutete darauf hin, daß tatsächlich beide Eltern ein defektes Allel des TT1 Gens tragen und vererbt haben.

5

Zur Klonierung eines Gens wurden DNA-Abschnitte bestimmt, die eine bekannte DNA-Sequenz, z.B. ein Transposon, flankieren. Dazu mußte zunächst gezeigt werden, daß sich in der Linie 5K60 das En-Transposon immer noch im TT1 Gen befand. Für diesen Zweck wurde eine Population von 51 Schwesterpflanzen der tt1-En Linie mittels Southern Blotting analysiert. 19 dieser Pflanzen produzierten gelbe und 33 dunkelbraune Samen. Es konnten verschiedene Banden identifiziert werden, die mit einer En-Sonde hybridisierten. Eine dieser Banden fand sich in allen Pflanzen, die gelbe Samen hervorbrachten, sowie in 16 der braunsamigen. Sie fehlte in den übrigen 17 Pflanzen. Dieses Ergebnis bestätigt die Annahme, daß alle gelbe Samen produzierende Pflanzen homozygot für eine Insertion des En-Transposon am tt1 Locus sind, während die braunsamigen Pflanzen entweder heterozygot für diese Insertion oder homozygot für den Wildtyp sind.

Die flankierende DNA dieser Insertion wurde durch schnelle Amplifikation genomischer Enden (RAGE), vgl. Cormack und Somssich, Gene 194, 273-276 (1997), gewonnen und in den pCR-TOPO Vektor (Firma Invitrogen) kloniert. Das resultierende Plasmid ist pCR-RAGE. Das Insert von pCR-RAGE wurde als Sonde verwendet, um die IGF-BAC Bank (Mozo et al., Plant Journal 16, 377-384 (1998)) zu durchsuchen. Dabei wurden 5 positive Klone identifiziert (11O6, 3N5, 2B22, 10P4, 4M12). Alle diese Klone sind im *Arabidopsis thaliana* Genom auf Chr.I bei etwa 55cM lokalisiert, was mit der Kartenposition der tt1-Mutation übereinstimmt.

25

Beispiel 4: Sequenzierung der genomischen TT1-Region

Ein 12 kb großes SpeI-Fragment des BACs 3N5, das mit der Sonde hybridisiert, wurde in pSK Bluescript (Firma Stratagene) subkloniert. Das resultierende Plasmid ist pSK-TT1. pSK-TT1 wurde unter Verwendung des Genome Priming Systems GPS1 der Firma New England Biolabs und eines Sequenzierautomaten der Firma ABI, Modell 377, durchsequenziert. Die Sequenzen wurden zusammengefügt und mit verschiedenen Computerprogrammen untersucht. Vergleiche von TT1 mit Sequenzen aus der NCBI GenBank ergaben Ähnlichkeiten mit Zinkfingerproteinen. Zinkfingerproteine sind in der Lage, an DNA-Sequenzen zu binden und regulatorische

30

Funktionen bezüglich der Expression bestimmter Gene auszuüben. Die von der TT1-cDNA abgeleitete Proteinsequenz von 303 AS Länge zeigt Ähnlichkeiten von etwas über 30% zu Zinkfingerproteinen wie StPCP1 (X82328, Kühn und Frommer 1995, MGG 247, 759-763) und ZmID1 (AF0058757, Colasanti et al, 1998, Cell 93, 593-603). Eine weitaus höhere Ähnlichkeit von über 70% besteht zu Datenbankeinträgen, die bislang allerdings nur hypothetische Proteine repräsentieren. Der Computervergleich zeigt neben der TT1-Aminosäuresequenz hypothetische Aminosäuresequenzen aus *Arabidopsis thaliana*, die aus den Sequenzen mit den Acc.No AL049660, AB025629 oder AC006085.9 und für *Hordeum vulgare* aus AJ2347041 abgeleitet wurden. Die Ähnlichkeiten erstrecken sich in diesem Fall über einen deutlich grösseren Bereich der Sequenz, gehen also über die Zinkfingerregion hinaus.

Beispiel 5: Ermittlung der TT1-cDNA

Die Existenz eines exprimierten Gens sowie die Position des putativen Introns wurden mittels RT-PCR überprüft. Durch 3' und 5'-Race wurde die Länge der vollständigen cDNA bestimmt.

Beispiel 6: Komplementation der tt1-Mutation

Um zu zeigen, daß es sich bei dem klonierten Gen tatsächlich um TT1 handelt, wurde die tt1 Mutation komplementiert. Dazu wurde das 12 kb Insert aus pSK-TT1 in die SpeI-Schnittstelle des Vektors pGPTV-Kan-TATA::GUS umklontiert. Dieser Vektor entstand aus pGPTV-Kan (Becker et al., Plant Molecular Biology 20, 1195-1197 (1992)) durch Austausch des CaMV 35S-Promotors gegen einen Polylinker. Nach Transformation in den Agrobakterienstamm GV3101 (mit dem Virulenzplasmid pMP90, Koncz und Schell, Molecular and General Genetics 204, 383-396 (1986)) wurden tt1-Pflanzen mittels Vakuuminfiltration transformiert (Bechthold et al., Molecular Biology and Genetics 316, 1194-1199 (1993)). Transformanten wurden auf kanamycinhaltigem MS-Medium selektiert und auf ihre Samenfarbe hin analysiert. Samen komplementierter Pflanzen entsprachen bezüglich ihrer Färbung dem Wildtyp.

Beispiel 7: Ermittlung der zellulären Lokalisation des TT1-Proteins

Um die zelluläre Lokalisation des TT1-Proteins zu ermitteln, wurde es c-terminal an das „Grün fluoreszierende Protein“ (GFP) fusioniert. Unter Verwendung von pAVA393 (von Arnim et al., Gene 221, 35-43 (1998)) und der kompletten cDNA entstand so pTT1-GFP. Das Plasmid wurde in *Arabidopsis*-Protoplasten transfiziert (Hartmann et al., Plant Molecular Biology 36, 741-54 (1998)) und nach zwanzigstündiger Inkubation die GFP-Fluoreszenz bestimmt. Anhand dieser

Untersuchung konnte die Lokalisation des TT1-Proteins im Zellkern nachgewiesen werden.

Beispiel 8: Expressionsanalysen

Zur Untersuchung des TT1-Promotors wurde ein 3 kb grosses Fragment zunächst in den Vektor pBT10 (Feldbrügge et al., Plant Journal 11, 1079-1093 (1997)) vor GUS gesetzt und die ganze Promotor-GUS Kasette danach in den binären Vektor pGPTV überführt. Nach Transformation in *Agrobacterium* wurden damit Wildtyp-Pflanzen von *Arabidopsis thaliana* Columbia infiltriert. Transformanten wurden auf kanamycinhaltigen Agarplatten selektiert. Nach etwa zehn Tagen wurden die Keimlinge auf Erde überführt und bis zur Samenreife im Gewächshaus unter Langtagbedingungen kultiviert. Die Aktivität des Reporterenzym β -Glucuronidase wurde über die Umsetzung von farblosem X-Gluc zu einem blauen Farbstoff nachgewiesen. Dazu wurde das Substrat in Keimlinge bzw. Blätter und Infloreszenzen mit Schoten verschiedener Entwicklungsstadien infiltriert und anschließend das Chlorophyll durch ethanolische Extraktion aus dem Gewebe entfernt (Figur 8).

Nach zweitägiger Inkubation in der Substratlösung war in den transgenen TT1-GUS-Pflanzen die Reportergenaktivität nachweisbar. Während im Wildtyp unter diesen Bedingungen keine Blaufärbung auftrat, war in den transgenen Pflanzen übereinstimmend GUS-Aktivität in geöffneten Blüten und sich entwickelnden Schoten zu beobachten. In geöffneten Blüten überragt das Gynoeceum die Antheren (Figur 8 A und B) und die Bestäubung hat bereits stattgefunden. Die Untersuchung des Pflanzenmaterials am Lichtmikroskop ergab, daß die GUS-Aktivität in den Funiculi und Integumenten der Samenanlagen sowie in weiteren mütterlichen Geweben (Scheidewand, Schotenbasis) nachweisbar war (Figur 8 C und D). Sie wurde vor allem in den obersten und den untersten Samenanlagen junger Schoten detektiert (Figur 8 B und E). Die GUS-Aktivität konnte in Schoten bis zum Stadium 17, nicht jedoch in älteren Stadien beobachtet werden.

5

Patentansprüche

10 1. Verfahren zur Herstellung einer Pflanze mit veränderter Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat, sofern das Fragment oder Derivat die Expression von Genen im Samen spezifisch steuert, und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz in das
15 Genom von Pflanzenzellen oder Pflanzengewebe und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

2. Verfahren nach Anspruch 1, wobei die Genexpression verstärkt oder verringert ist.

20 3. Verfahren nach Anspruch 1 oder 2, wobei für die für ein Genprodukt codierende Nukleinsäuresequenz eine endogene oder exogene Nukleinsäuresequenz verwendet wird.

4. Verfahren nach einem der Ansprüche 1 bis 3, wobei für die für ein Genprodukt codierende Nukleinsäuresequenz ausgewählt aus der Gruppe der Gene des
25 Phenylpropanoidstoffwechsels, samenspezifischer Gene, samenschalenspezifischer oder der Gene des allgemeinen Stoffwechsels verwendet wird.

5. Verfahren nach Anspruch 4, wobei für die Gene des Phenylpropanoidstoffwechsels eine Nukleinsäuresequenz verwendet wird ausgewählt aus der Gruppe der Gene für Phenylalanin-
30 Ammonium-Lyase, Cinnamat-4-Hydroxylase, 4-Coumarat-CoA-Ligase, Chalcon-Synthase, Chalcon-Isomerase, Chalcon-Reduktase, Flavanon 3-Hydroxylase, Flavonoid-3'-Hydroxylase, Flavonoid-3'5'-Hydroxylase, Dihydroflavonol-4-Reduktase, Leucoanthocyanidin-Reduktase, Leucoanthocyanidin-Dioxygenase, 3'-Glucosyltransferase, 5'-Glucosyltransferase und O-Methyl-Transferase.

35

6. Verfahren nach Anspruch 4, wobei für die samenschalenspezifischen Gene eine Nukleinsäuresequenz verwendet wird ausgewählt aus der Gruppe der die Keimstimmung oder Dormanz beeinflussenden Gene, der die Pathogenresistenz beeinflussenden Gene und des TT1-Gens gemäß SEQ ID NO:2 und SEQ ID NO:4.

5

7. Verfahren nach Anspruch 4, wobei für die Gene des allgemeinen Stoffwechsels eine Nukleinsäuresequenz verwendet wird ausgewählt aus der Gruppe der Gene für ADP-Glucose-Synthetase, Stärkesynthase, ADP-Glucose-Pyrophosphorylase und Hefe-Invertase.

10 8. Verfahren nach einem der Ansprüche 1 bis 7, wobei für die samenspezifische regulatorische Sequenz die Nukleinsäuresequenz gemäß SEQ ID NO:1 oder deren Fragment oder Derivat verwendet wird.

15 9. Transformierte Pflanzenzelle oder transformiertes Pflanzengewebe, dadurch gekennzeichnet, daß eine samenspezifische regulatorische Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz stabil in das Genom der Pflanzenzelle oder des Pflanzengewebes integriert ist.

20 10. Nukleinsäuresequenz gemäß SEQ ID NO:1.

11. Fragment oder Derivat der Nukleinsäuresequenz nach Anspruch 10 oder eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert und für die samenspezifische Expression verantwortlich ist.

25

12. Nukleinsäuresequenz nach Anspruch 11, wobei die hybridisierende Nukleinsäuresequenz unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert.

13. Verfahren zur Herstellung von Pflanzen mit verändertem Flavonoidgehalt, umfassend
30 das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat mit der biologischen Aktivität eines Polypeptids codiert durch die Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der

erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

14. Verfahren nach Anspruch 13, wobei die integrierte Nukleinsäuresequenz oder deren Fragment oder Derivat in Sense- oder Antisense-Orientierung zur den Flavonoidgehalt steuernden endogenen Nukleinsäuresequenz exprimiert wird.
15. Verfahren nach Anspruch 13 oder 14, wobei die Bildung von Flavonoiden durch ein Ribozym, umfassend die integrierte Nukleinsäuresequenz oder deren Fragment oder Derivat unterdrückt wird.
16. Verfahren nach Anspruch 13 oder 14, wobei die Nukleinsäuresequenz oder deren Fragment oder Derivat in den genomischen Bereich des homologen endogenen Gens durch homologe Rekombination integriert wird.
17. Verfahren nach einem der Ansprüche 13 bis 16, wobei die Nukleinsäuresequenz oder deren Fragment oder Derivat funktional mit einer regulatorischen DNA-Sequenz verbunden ist, die die Expression der integrierter Nukleinsäuresequenz oder deren Fragment oder Derivat steuert.
18. Verfahren nach Anspruch 17, wobei die regulatorische DNA-Sequenz ausgewählt ist aus der Gruppe der Promotoren CaMV 35S-Promotor, PRPI-Promotor, Phaseolin-Promotor, Isoflavon-Reduktase Promotor, ST-LSI Promotor, durch Salizylsäure induzierbarer Promotor, durch Benzensulfonamid induzierbarer Promotor, durch Tetrazyklin induzierbarer Promotor, durch Abscisinsäure induzierbarer Promotor, durch Ethanol oder Cyclohexanon induzierbarer Promotor, Promotor gemäß SEQ ID NO:1 oder ein samenspezifischer Promotor aus Tabak.
19. Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4.
20. Fragment oder Derivat der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4, oder eine Nukleinsäuresequenz die mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert und für die Bildung von Flavonoiden verantwortlich ist.

21. Nukleinsäuresequenz nach Anspruch 20, wobei die hybridisierende Nukleinsäuresequenz unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert.
- 5 22. Transformierte Pflanzenzelle oder transformiertes Pflanzengewebe, dadurch gekennzeichnet, daß die Nukleinsäuresequenz nach Anspruch 19 bis 21 stabil in das Genom der Pflanzenzelle oder des Pflanzengewebes integriert ist.
23. Aminosäuresequenz wie in SEQ ID NO:3 aufgeführt.
- 10 24. Pflanzenzelle oder Pflanzengewebe nach Anspruch 9 oder 22, regenerierbar zu einer samenproduzierenden Pflanze.
25. Pflanze, erhältlich nach einem der Ansprüche 1 bis 8 oder 13 bis 18.
- 15 26. Saatgut, erhalten von Pflanzen nach Anspruch 25.
27. Vektor, umfassend eine Nukleinsäuresequenz nach einem der Ansprüche 10 bis 12 oder 19 bis 21.
- 20 28. Transgene Pflanze mit einer stabil in das Genom integrierten samenspezifischen regulatorischen Nukleinsäuresequenz gemäß SEQ ID NO:1, oder deren Fragment oder Derivat oder Homolog mit der biologischen Funktion eines samenspezifischen Promotors, und einer mit dieser Nukleinsäuresequenz funktionell verbundenen für ein Genprodukt codierenden
- 25 Nukleinsäuresequenz.
29. Transgene Pflanze mit einer stabil in das Genom integrierten Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat mit der biologischen Aktivität eines Polypeptids codiert durch die
- 30 Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4.
30. Transgene Pflanze nach Anspruch 29, wobei die Nukleinsäuresequenz oder deren Fragment oder Derivat funktional mit einer regulatorischen DNA-Sequenz verbunden ist, die die Expression der integrierter Nukleinsäuresequenz oder deren Fragment oder Derivat steuert.

1 actagttgaccacatgaactaaacttcttggacaatcatcaatggacaca
51 tgttagcctttgatttgctgtgaatttggtttatctctcagtataattatc
101 actttcttggtttatgcttacaatatattttatgggttagagttttgtttt
151 acgattttggatttaatggataaagattagggattgagggtttgagttta
201 gggtaaggaaattaggcttttagtgtagagtcctcaagggtttaaggtttac
251 acaccacaaaccatttgcttgtgtcaacaacattgtatcatattttcaaa
301 aaaattttggtgaaggaccttgattgatataataaagcgaactgtttg
351 gataagtttatgtggacaatatattggatacataattagaaacatagt
401 ttaatatctgatatttggtgggaatatataatactacttaggtttaata
451 tatagtattcatatgatgcgaactgtttggataagtttacgtggacaat
501 atatatgtgatacataattaggaacatagtttaatatgtatattgttg
551 ggaatatataattctacttacgcttaaatatttttattgaattaaagca
601 tttcatataatgtgaactgtttgaatatgtttacatggacaatatatatt
651 ggatacataattaggaacatagtttaatatctgatatttggtggaatat
701 ataatatagtttaagcttaaatatttttatttgatataatattgactta
751 aacattttttatttgattaaactaaattttaacagatcttaccattaattt
801 ttaacttggtatctctatctaagtgcacgtatattgttttttagtaattg
851 gcaacaaaattaatttatctctgttttttttctctcacctttataag
901 ggtaaaatgggtcataaaatcagtaaaaaagggtggaagtgccactccc
951 tcaaaagtggtcataaacgtccaaactttctccataaatgccttattttgg
1001 aacattccatataagattataacttattataggttataacttattatagtt
1051 acgttaattatatgaatttctatttagttatcacacaatcaaaTATTTTAA
1101 TCACAAaaattttattaaacattttatatgtggtagtataatgcaataaca
1151 tattatatgtggtggcataatgcaacaacatatattttgtctacgaatct
1201 cctttatttttcgtttatgtaacaacagtaaaacggattgtttagcttga
1251 tattctatattataataatctaaagttatttttgtaaattattttttttc
1301 caaattggataaccaatcatagagttggtatttttttTTTGTGTTAAAT
1351 Atatatatactgaatatcgagttattgtgcatgacaatttatatggcgga
1401 cgagtttaaatcgacattaataacaattaaaatattattaatctaatact
1451 taaatactgggttaaataccaatttattattcttaataccacatatataa
1501 catatctaactcttactgattcaataaagattgtgtgaaacaaaagttgt
1551 cttgcaaagaattaatattgtacatagattttgttctggtagctagtact
1601 aaaatccattaataaaaactaatacgggtatctttattgatcatgtaacatg
1651 aattattcatgtatatacaattgaccctattaattttgcataaaattcaa
1701 cttggcaaattcattgattttgtaaacggttaattctgctaatttcacaa
1751 ttctctgtacgctaaaaatttatgcgtattatcgattgatatgcaaat
1801 atcgaagaatttatagttttatatagtagaaatgaaggattttgcaaaac
1851 gagttctaacgtgaaataacactaattaattaattagagtttgaaacctac
1901 agagattcgacttgatccacttgaaaaattcatttactctactaatttgg
1951 ttactccatggaccatgattatgctattctgtaggactctaacaactgac
2001 ttgacacaatctcttctgtgaacaataatgggttatattttttggtttg
2051 ttttttcggacaaattagccacgttgcttttagaccattttgtagttctta

2101 tcttgaatcaaagtctcagctaaaaaaaaAAAAACGCTTAAatccac
2151 tagctagactacgactacggttggttaaatgttttTTTTTAAATACAAtac
2201 aTTGAAGTTAAATATttgaataaagaaaatctaatacagcatgtatacagt
2251 atattagaagtaatacttgatcagaaaaataacatacaataataaaataa
2301 taaaaaaattatgttagtttttgggaatattataattctactttcaatca
2351 aaataactaaaagaaataaaatcttcacacatagtGGTAATAATTGGCTa
2401 gtatgaatattgaattgtggagaccggcataatatttgactaggcagaa
2451 attattgatattgactaagttaataaccttgcaaagaaattcttttagtg
2501 aaacgtgtacatttgtaaaaacagatttaacactaaatcttgacttgtat
2551 atactattaattatctcttttctcttattgggtatgtcaaactctagtgtt
2601 acaaaaccagagggtgttgaccgttagagagagaattaaacaacttacata
2651 catacaaaacataacccaaaaaataataataatgcatcttccataataa
2701 taataatatgaattcaacattagcatttcatttcattacccaaatccgaaa
2751 tttcattgattaaaattaatacaattgtattgtagaaaagctaaaagctt
2801 acgtaaatgccaaagatagtcaaaaccctgcaatgacaaaagttgccaaaa
2851 tcttgaagagtttgggtccacaaaatttaaggttcttggttttccactcta
2901 tttataggcaaagagatgagacagagaagattaaattacttcttaacaaa
2951 gggtgttttctactcaaccacatgcattctcaagtgctgctcctcacatt
3001 ccccaagattcccatttactcacttctctatttggtacgtaagtcacaca
3051 atatgattctaaattattttacacattattcggtttgttcacacttgctt
3101 tcgactttcgtaaacctatattagttcatccaatattattcggtaaattcg
3151 atatttatcaatctttattctcgtagggttaaaggagacgattgatacgtg
3201 ggatctacttacgtatctgcatgattattagttataaaaagttattgcaa
3251 cattaaattactttcatagagagcaatcattatattaaggtaatttaatt
3301 ttattatatatagtcaagatttaaggaataaagaaaagattctcaaac
3351 atttcattctctctccaacaactattcaccacattcaATG

3387 ATGGAGTCACCACC
3401 ACTATACGAGATATCCTCAAGCTCTTCTTCTGAAAAACCTAGACACCATT
3451 TCCAATCCCTTGATCTCTTCCCTAACCTCAACCAAACTCTTGTATCAAC
3501 AATACCCTAATTGAGCCTTTACCGCTTATTGATCGCATAACTTGAAGTC
3551 AAACCTAGACCTAAACCCTAATCCCTTGTATGCGGAAGAAGGAGAGCAAG
3601 AGGAGGAAGAAGAAGAAGAAGACCGTGAAGTGGACGTGGACTTACAC
3651 ATCGGCCCTTCCTGGTTTTGGTAAACCAAGCAATGATGCTAAACAGCTGAA
3701 GAAGAGAAATGGGAAGGAGATCGCCACATATGACGCCGAAAAGGCATCG
3751 AGAATGAACCTTCCGGAAGGCATACTGGATCCCGGCGCCGGAGCAAATT
3801 CTCATAGGGTCACTCATTTTCTTGCCATGTATGCTTCAAGACATTCAA
3851 TCGCTACAACAATCTTCAGGtacgagtcaatataatctcatgcgcatgtgt
3901 tttccatgcacaaacatatataataaattcatcttatagagttatatctc
3951 cggatctaattgttatgagtttatctatatctatatataacatatata
4001 tatatatatatatatatatataataattctgaatttatgtgataaaagc
4051 taaacaaccaggatttaataagatgatttaccttggatcttattatacaa
4101 tttacaaatttaatacaagtcaactaatcgtgatttaattacttttttttg
4151 taagaagagttggtaatatataatttttatggtaattgtttcatgaaaaata
4201 attcatcacaaactctttacattttatattaatgccttaactaaagctgaatt
4251 cgaaaaagttgaaataaattatctactaagatttgattgactatagtttt
4301 taatagttttcttttctcatatatataattatcatagtagtcaaaacattt
4351 gattcaaaacttaaatcacagatttcttgaaatgaaacattactatgctcg
4401 gtcaataatatgatttttaaggaaccatgttatcttcttttattacttaag
4451 gaaacctttttgttttttgggtgactctaaatattatgaaatatagATGCAC
4501 ATGTGGGGACATGGTTCACAATACAGGAAAGGACCGGAGTCACTGAAAGG
4551 CACACAGCCACGAGCCATGTTAGGGATCCCTTGTTACTGCTGCGTTGAAG
4601 GGTGCAGGAACCACTTGACCATCCTCGTTCGAAGCCACTGAAAGACTTT
4651 AGGACGCTCCAAACGCACTACAAACGCAACACGGACACAAACCCTTCTC
4701 GTGTCGCCTTTGCGGTAAGCTTTTGGCTGTCAAGGGCGATTGGCGAACAC
4751 ATGAGAAGAATTGTGGAACGTTGGGTTTGCCTTTGCGGTTCTGATTTT
4801 AAACACAAACGTTCTCTTAAGGACCATGTTAAGGCGTTGGGTCTGGTCA
4851 TGGGCCTTATCCAACCTGGTTTGTGTTGAAGAGCAGGCTTCTAATTCATCTG
4901 TCTCCGAGACTTTGTTTTTTTAAatttgggcatcttttctttcgttat
4951 gaaatatctatttacttttagaaaaataaatgtggtatctaattgttcc
5001 aaattaggaacacgaagtgtaccattatatttttcatcactacaaatgtt
5051 attcagagaaaattatcatattgtctcgttaaagatagaatagggttt
5101 gaatttatcaaatattaaaaacagatcaatacaaaattgaccatgcatat
5151 gcacttgaatattctgatttctttatgatgtaatctcattcaagaaaagc
5201 tt

FIG: 3

ATGGAGTCACCACCACTATACGAGATATCCTCAAGCTCTTCTTCTGAAAAACCTAGACAC
1 -----+-----+-----+-----+-----+-----+-----+ 60
M E S P P L Y E I S S S S S S E K P R H -

CATTTCCAATCCCTTGATCTCTTCCCTAACCTCAACCAAACTCTTGTATCAACAATACC
61 -----+-----+-----+-----+-----+-----+-----+ 120
H F Q S L D L F P N L N Q N S C I N N T -

CTAATTGAGCCTTTACCGCTTATTGATCGCATAACTTGAACCTAAACCTAGACCTAAAC
121 -----+-----+-----+-----+-----+-----+-----+ 180
L I E P L P L I D R I N L N S N L D L N -

CCTAATCCCTTGTATGCGGAAGAAGGAGAGCAAGAGGAGGAAGAAGAAGAAGAAGAC
181 -----+-----+-----+-----+-----+-----+-----+ 240
P N P L Y A E E G E Q E E E E E E E E D -

CGTGAAGTGGACGTGGACTTACACATCGGCCTTCCTGGTTTTGGTAAACCAAGCAATGAT
241 -----+-----+-----+-----+-----+-----+-----+ 300
R E V D V D L H I G L P G F G K P S N D -

GCTAAACAGCTGAAGAAGAGAAATGGGAAGGAGATCGCCACATATGACCCGGAAAAGGC
301 -----+-----+-----+-----+-----+-----+-----+ 360
A K Q L K K R N G K E I A T Y D A G K G -

ATCGAGAATGAACCTTCCGGAAGGCATACTGGATCCCGCGCCGAGCAAATTCTCATA
361 -----+-----+-----+-----+-----+-----+-----+ 420
I E N E L S G K A Y W I P A P E Q I L I -

GGGTCACTCATTCTTCTTGCCATGTATGCTTCAAGACATTCAATCGCTACAACAATCTT
421 -----+-----+-----+-----+-----+-----+-----+ 480
G F T H F S C H V C F K T F N R Y N N L -

CAGATGCACATGTGGGGACATGGTTCACAATACAGGAAAGGACCGAGTCACTGAAAGGC
481 -----+-----+-----+-----+-----+-----+-----+ 540
Q M H M W G H G S Q Y R K G P E S L K G -

ACACAGCCACGAGCCATGTTAGGGATCCCTTGTACTGCTGCGTTGAAGGGTGCAGGAAC
541 -----+-----+-----+-----+-----+-----+-----+ 600
T Q P R A M L G I P C Y C C V E G C R N -

CACATTGACCATCCTCGTTCCAAGCCACTGAAAGACTTTAGGACGCTCCAAACGCACTAC
601 -----+-----+-----+-----+-----+-----+-----+ 660
H I D H P R S K P L K D F R T L Q T H Y -

AAACGCAAACACGGACACAAACCTTCTCGTGTGCGCTTTGCGGTAAGCTTTGGCTGTC
661 -----+-----+-----+-----+-----+-----+-----+ 720
K R K H G H K P F S C R L C G K L L A V -

AAGGGCGATTGGCGAACACATGAGAAGAATTGTGGAACCGTTGGGTTTGGCTTTGCGGT
721 -----+-----+-----+-----+-----+-----+-----+ 780
K G D W R T H E K N C G K R W V C V C G -

TCTGATTTTAAACACAAACGTTCTCTTAAGGACCATGTTAAGGCGTTTGGGTCTGGTCAT
781 -----+-----+-----+-----+-----+-----+-----+ 840
S D F K H K R S L K D H V K A F G S G H -

GGGCCTTATCCAACCTGGTTTGTGTTGAAGAGCAGGCTTCTAATTCATCTGTCTCCGAGACT
841 -----+-----+-----+-----+-----+-----+-----+ 900
G P Y P T G L F E E Q A S N S S V S E T -

TGTTTTTTTAA
901 -----+----- 912
L F F * -

	1		50
AtTT1
AtAL049660
AtAB025629	MLFSTVLSHR	TLYILTCPNT	LIHSYTHPHI HAYLAFTGFL TQLHHLEISC
AtAC006085.9
Hv234704
Consensus
	51		100
AtTT1MESPP LYEISSSSSS
AtAL049660MTDP YSNFFTDWFK SNPFHH..YP NSSTNPSHP LPPVTPPSSF
AtAB025629	LLLLFFSLSS	LLKLMADPDC	IFRNGYVDYY NYSFNYATSL SRIYNSHDSF
AtAC006085.9MSNPAC SNLFNNGCDH N.SFNYSTSL SYIYNSHGSY
Hv234704
ConsensusS.
	101		150
AtTT1	EKPRHHFQSL	DLFPNLNQNS	CINNTLIEPL PLIDRINLNS NLDLNPNP..
AtAL049660	FFFPQSGD..	LRRPPPPPTP	PPSPPLREAL PLLSLSPANK QQDHHNH.D
AtAB025629	YYPHQTTPN	INE.NPNLTS	PDSPLREAL PLLSLSPIHK HQEPTANHHE
AtAC006085.9	YYSNTTNPY	INHHTTSTS	PNSPLREAL PLLSLSPI.R HQEQDQH..
Hv234704
Consensusl.e.l pl.....
	151		200
AtTT1	.LYAEEGEQE	EEEEEEEDREVDVDLHIG LPGFG.....
AtAL049660	HLIQEPPSTS	MDVDYDHHQ	DDHHNLDDDD HDVTVALHIG LPSPSAQEMA
AtAB025629	YYFMETTETS	SNSNFLDQCQ	DSYR.....HDVTVDLHLG LPNLGDGG..
AtAC006085.9	.YFMDTHQIS	S.SNFLDDPLVTVDLHLG LPNYGVGE..
Hv234704
Consensusv.v.lh.g lp.....
	201		250
AtTT1KPSND	AKQLKCRNGK	EIATYDAGKG IENELSGKA.
AtAL049660	SLMMSSSSSS	SSRTHHHED	MNHKKDLDE YSHGAVGGGE DDDSDSVGGD
AtAB025629SSSSD	VULDSTDHQE	GHHDHHDQDG LEVTMAS... DHDDHGGGLQ
AtAC006085.9SIRS	IAPDATTDEQ	...DQDHRG VEVTVESHLD DDDHHHGLD
Hv234704
Consensus
	251		300
AtTT1YWIPAPEQ	ILIGFTFSC HVCFKTFNRY NNLQMHMWGH
AtAL049660	GGCRISRLNK	GQYWIPTPSQ	ILIGPTQFSC PVCFKTFNRY NNMQMHWGH
AtAB025629	RGNHLLH...	..FWIPTPSQ	ILMGPTQFSC PLCFKTFNRY NNMQMHWGH
AtAC006085.9	RG...HH...	..YWIPTPSQ	ILIGPTQFSC PLCFKTFNRY NNMQMHWGH
Hv234704QMHWGH
Consensuswip.p.q	il.g.t.f.c ..cfktnry nn.QMHWGH
	301		350
AtTT1	GSQYRKGPES	LKGTQPRAML	GIPCYCCVEG CRNHIDHPRS KPLKDFRTLQ
AtAL049660	GSQYRKGPES	LRGTQPTGML	RLPCYCCAPG CRNNIDHPRA KPLKDFRTLQ
AtAB025629	GSQYRKGPES	LRGTQPTAML	KLPCYCCAPG CKNNIDHPRA RPLKDFRTLQ
AtAC006085.9	GSQYRKGPES	LRGTQPTGML	RLPCYCCAPG CKNNIDHPRA KPLKDFRTLQ
Hv234704	GREYRKGPES	LKGTQTVALL	KVPCYCAA.G CRNSVSHPR RPLKDFRT..
Consensus	Gs#YRKGPES	LkGTQp.a\$L	k.PCYCea.G CrN.!dHPra rPLKDFRTLq
	351		400
AtTT1	THYKRKHGK	PFSCRLCGKL	LAVKGDWRTH EKNCGRWVC VCGSDFKHKR
AtAL049660	THYKRKHGK	PFMCRKCGKA	FAVRGDWRTH EKNCGLWYC ICGSDFKHKR
AtAB025629	THYKRKHGVR	PFACRRCGKA	FAVKGDWRTH EKNCGLWYC SCGSDFKHKR
AtAC006085.9	THYKRKHGSK	PFACRMCGKA	FAVKGDWRTH EKNCGLWYC SCGSDFKHKR
Hv234704
Consensus	thykrkhg..	pf.cr.cgk.	.av.gdwrtth ekncgk.w.c .cgsdfkhkr
	401		448
AtTT1	SLKDHVKAFG	SGHGYPYPTG.	.LFEEQASNS SVSETLFF..
AtAL049660	SLKDHVKAFG	NGHGAYGID.	.GFDEED..E PASEVEQLDN DHESMQSK
AtAB025629	SLKDHVKAFG	NGHVPC....	CGIDHEEE.E AASDVEQQE.
AtAC006085.9	SLKDHVKAFG	NGHVPCGIDS	FGGDHEDYYD AASDIEQ... ..
Hv234704
Consensus	slkdh.kafg	.gh.....s.....

FIG. 5

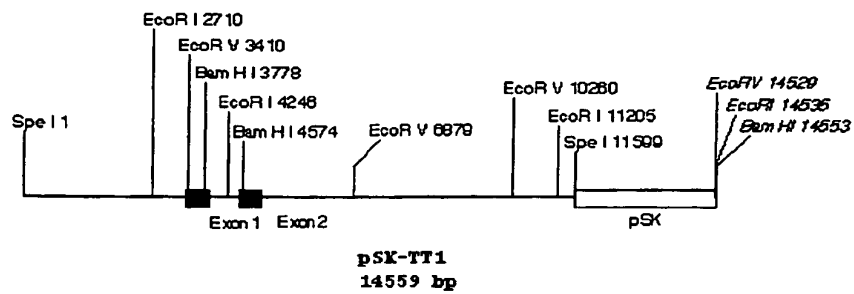


FIG. 6

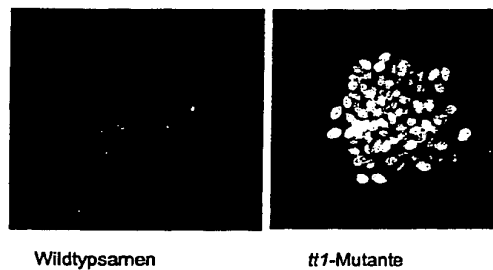


FIG. 7

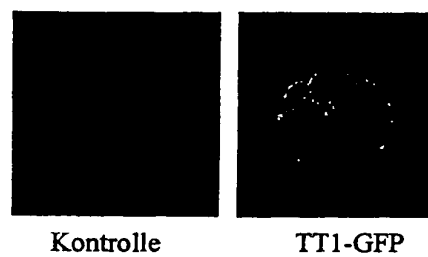
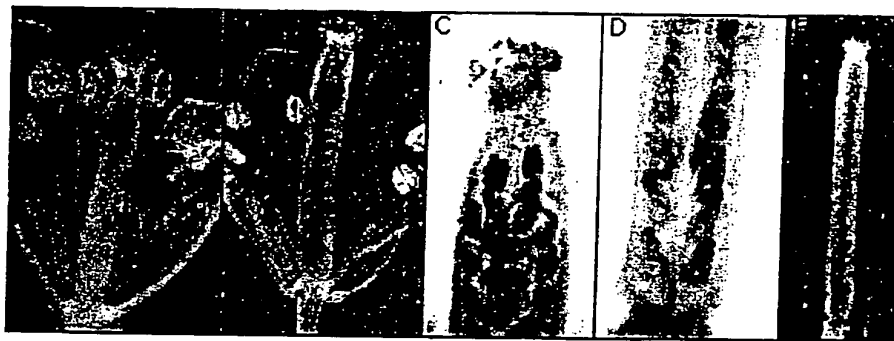


FIG. 8



SEQUENZPROTOKOLL

<110> Max-Planck-Gesellschaft z. Förd. d. Wissenschaften

5 <120> Pflanzen mit veränderter Genexpression

<130> GI-001

<140> xx

10 <141> 1999-07-02

<160> 4

<170> PatentIn Ver. 2.1

15

<210> 1

<211> 3389

<212> DNA

<213> *Arabidopsis thaliana*

20

<400> 1

actagttgac cacatgaact aaacttcttg gacaatcatc aatggacaca tgtagctttt 60
gatttgctgt gaatttggtt tatctctcag tataattatc actttcttgt ttatgcttac 120
aatataatatt atgggttaga gttttgtttt acgatttttg atttaattga taaagattag 180
25 ggattgaggg tttgagttta gggttaaggaa attaggcttt agtgtagagt ctcaaggggt 240
taagggtttac acaccacaaa ccatttgctt gtgtcaacaa cattgtatca ttttttcaaa 300
aaaattttgt tgaaggacct tgtattgata tatataaagc gaactgtttg gataagttta 360
tgtggacaat atatatgtga tacataatta gaaacatagt ttaatatctg atatttggtg 420
ggaatatata atactactta ggtttaaata tatagtattt catatgatgc gaactgtttg 480
30 gataagttta cgtggacaat atatatgtga tacataatta ggaacatagt ttaatatctg 540
atatttggtg ggaatatata attctactta cgcttaaaata tttttatttg aattaaagca 600
tttcatataa tgtgaactgt ttgaatatgt ttacatggac aatataatatt ggatacataa 660
ttaggaacat agtttaatat ctgatatattg ttggaaatat ataatatag ttaagcttaa 720
atatttttat ttgatataat atttgactta aacattttta tttgattaaa cttaaattta 780
35 acagatctta ccattaattt ttaacttggt atctctatct aatgtcacgt atattgtttt 840
ttagtaattg gcaacaaaat taatttatct cctgtttttt ttccttctca cttttataag 900
ggtaaaatgg tcataaaatc agtaaaaaag gtggaaaagt gccactccc tcaaaagtgt 960
cataaacgct caaactttct ccataaatgc cttatttttg aacattccat atagattata 1020
acttattata ggttataact tattatagtt acgttaatta tatgaatttc tattagtatt 1080
40 cacacaatca aatattttta tcacaaaaat ttattaaaca ttttatatgt ggtagtataa 1140
tgcaataaca tattatatgt ggtggcataa tgcaacaaca tattatttgt ctacgaatct 1200
cctttatttt tcgtttatgt aacaacagta aaacggattg tttagcttga tattctatat 1260
tataataatc taaagttatt tttgtaaatt attttttttc caaattggat aaccaatcat 1320
agagttggta tttttttttt ttgttaaaat atatatatac tgaatatcga gttattgtgc 1380
45 atgacaattt atatggcgga cgagtttaaa tcgacattaa taacaattaa aatattatta 1440
atctaatact taaatactgg ttaaatcacc aattttattt tcttaatacc acatattaaa 1500
catatctaatt ctttactgat tcaataaaga ttgtgtgaaa caaaagttgt cttgcaaaga 1560
attaatatgt tacatagatt ttgttctggt agctagtact aaaatccatt aataaaacta 1620
atacggatc tttattgatc atgtaacatg aattattcat gtatatacaa ttgaccttat 1680
50 taattttgca taaattttca cttggcaaat tcattgattt tgtaaaccgt taattctgct 1740
aatttcacaa ttctcttgta cgctaaaaat ttatgcgtat tatcgattg atatgcaaat 1800
atcgaagaat ttatagtttt atatatagta aatgaaggta tttgcaaaac gagttcctaac 1860
gtgaaataac actaattaat taattagagt ttgaacctac agagattcga cttgatccac 1920
ttgaaaaatt catttactct actaatgttg ttactccatg gacctgatt atgctattct 1980
55 gtaggactct aacaactgac ttgacacaat ctctttcgtg aacaataatg ggttatattt 2040
ttttgttttg ttttttcgga caaattagcc acgttgcttt agaccatttt gtagttctta 2100
tcttgaatca agtctcagc taaaaaaaaa aaaaaaacgc ttaaatccac tagctagact 2160
acgactacgt tgggttaaatg ttttttttta aatacaatac attgaagtta aatatttgaa 2220
taaagaaaaa ctaatacagc tgtatacagt atattagaag taatacttga tcagaaaaat 2280
60 aacatacaat aataaaaaaa taatacaagt atgttagttt ttgggaatat tataattcta 2340
ctttcaatca aaataactaa aagaaataaa atcttcacac atagtggtaa taattggcta 2400

gstatgaatat tgaattgtgg agacccggca taatatttga ctaggcagaa attattgata 2460
 tgtactaagt taataacctt gcaaagaaat tcttttagtg aaacgtgtac atttgtaaaa 2520
 acagatttaa cactaaatct tgacttgtat atactattaa ttattccttt tctcttattg 2580
 gtagtcaaaa tctagtgttt acaaaaccag aggtgttgac cgtagagag agaattaaac 2640
 5 aacttacata catacaaaac ataacccaaa aaaataataa taatgcatct tccataataa 2700
 taataatatg aattcaacat tagcattcat ttcatatccc aaatccgaaa tttcattgat 2760
 taaaattaat acaattgtat tgtagaaaag ctaaaagctt acgtaaatgc caaagatagt 2820
 caaaaccctg caatgacaaa gttgccaaaa tcttgaagag tttgggtccac aaaatttaag 2880
 gttcttgttt ttccactcta tttataggca aagagatgag acagagaaga ttaaattact 2940
 10 tcttaacaaa gggtgttttc actcaaccac atgcattctc aagtgtctgc tcctcacatt 3000
 ccccaagatt cccatttact cacttctcta tttggtacgt aagtcacaca atatgattct 3060
 aaattatttt acacattatt cgttttgttc acacttgctt tcgactttcg taaacctata 3120
 tagttcatcc aatattattc ggtaaattcg atatttatca atctttattc tcgtagggtta 3180
 aaggagacga ttgatacgtg ggatctactt acgtatctgc atgattatta gttataaaaag 3240
 15 ttattgcaaa cattaaatta ctttcataga gagcaatcat tatattaagg taattttaatt 3300
 ttattatata tagtcaagat ttaaaggaat aaagaaaaga ttctcaaaac atttcatctc 3360
 tctccaacaa ctattcacca cattcaatg 3389

20 <210> 2
 <211> 912
 <212> DNA
 <213> *Arabidopsis thaliana*

25 <400> 2
 atggagtcac caccactata cgagatatcc tcaagctctt cttctgaaaa acctagacac 60
 catttccaat cccttgatct cttccctaac ctcaaccaa actcttgat caacaatacc 120
 ctaattgagc ctttaccgct tattgatcgc ataaacttga actcaaacct agacctaaac 180
 30 cctaatecct tgtatgcgga agaaggagag caagaggagg aagaagaaga agaagaagac 240
 cgtgaagtgg acgtggactt acacatcggc cttcctgggt ttggtaaaacc aagcaatgat 300
 gctaaacagc tgaagaagag aaatgggaag gagatcgcca catatgacgc cggaaaaggc 360
 atcgagaatg aactttccgg aaaggcatac tggatccgg cgccggagca aattctcata 420
 gggttcactc atttttcttg ccatgtatgc ttcaagacat tcaatcgcta caacaatctt 480
 cagatgcaca tgtggggaca tggttcaca tacaggaaag gaccggagtc actgaaaggc 540
 35 acacagccac gagccatggt agggatccct tgttactgct gcgttgaagg gtgcaggaaac 600
 cacattgacc atcctcgttc caagccactg aaagacttta ggacgctcca aacgcactac 660
 aaacgcaaac acggacacaa acccttctcg tgtcgcttt gcggtaagct tttggctgtc 720
 aagggcgatt ggcgaacaca tgagaagaat tgtggaaaac gttgggtttg cgtttgcggt 780
 tctgatttta aacacaaaacg tttctttaag gaccatgtta aggcgttttg gtctgtgtcat 840
 40 gggccttatc caactggtt gtttgaagag caggcttcta attcatctgt ctccgagact 900
 ttgttttttt aa 912

<210> 3
 45 <211> 303
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 3
 50 Met Glu Ser Pro Pro Leu Tyr Glu Ile Ser Ser Ser Ser Ser Ser Glu
 1 5 10 15
 Lys Pro Arg His His Phe Gln Ser Leu Asp Leu Phe Pro Asn Leu Asn
 20 25 30
 55 Gln Asn Ser Cys Ile Asn Asn Thr Leu Ile Glu Pro Leu Pro Leu Ile
 35 40 45
 Asp Arg Ile Asn Leu Asn Ser Asn Leu Asp Leu Asn Pro Asn Pro Leu
 60 50 55 60

Tyr Ala Glu Glu Gly Glu Gln Glu Glu Glu Glu Glu Glu Glu Asp
 65 70 75 80
 5 Arg Glu Val Asp Val Asp Leu His Ile Gly Leu Pro Gly Phe Gly Lys
 85 90 95
 Pro Ser Asn Asp Ala Lys Gln Leu Lys Lys Arg Asn Gly Lys Glu Ile
 100 105 110
 10 Ala Thr Tyr Asp Ala Gly Lys Gly Ile Glu Asn Glu Leu Ser Gly Lys
 115 120 125
 Ala Tyr Trp Ile Pro Ala Pro Glu Gln Ile Leu Ile Gly Phe Thr His
 130 135 140
 15 Phe Ser Cys His Val Cys Phe Lys Thr Phe Asn Arg Tyr Asn Asn Leu
 145 150 155 160
 20 Gln Met His Met Trp Gly His Gly Ser Gln Tyr Arg Lys Gly Pro Glu
 165 170 175
 Ser Leu Lys Gly Thr Gln Pro Arg Ala Met Leu Gly Ile Pro Cys Tyr
 180 185 190
 25 Cys Cys Val Glu Gly Cys Arg Asn His Ile Asp His Pro Arg Ser Lys
 195 200 205
 Pro Leu Lys Asp Phe Arg Thr Leu Gln Thr His Tyr Lys Arg Lys His
 210 215 220
 30 Gly His Lys Pro Phe Ser Cys Arg Leu Cys Gly Lys Leu Leu Ala Val
 225 230 235 240
 35 Lys Gly Asp Trp Arg Thr His Glu Lys Asn Cys Gly Lys Arg Trp Val
 245 250 255
 Cys Val Cys Gly Ser Asp Phe Lys His Lys Arg Ser Leu Lys Asp His
 260 265 270
 40 Val Lys Ala Phe Gly Ser Gly His Gly Pro Tyr Pro Thr Gly Leu Phe
 275 280 285
 Glu Glu Gln Ala Ser Asn Ser Ser Val Ser Glu Thr Leu Phe Phe
 290 295 300
 45
 <210> 4
 <211> 1816
 50 <212> DNA
 <213> *Arabidopsis thaliana*
 <400> 4
 atggagtcac caccactata cgagatatcc tcaagctctt cttctgaaaa acctagacac 60
 55 catttccaat cccttgatct cttccctaac ctcaaccaa actcttgat caacaatacc 120
 ctaattgagc ctttaccgct tattgatcgc ataaacttga actcaaacct agacctaaac 180
 cctaatacctt tgtatgcgga agaaggagag caagaggagg aagaagaaga agaagaagac 240
 cgtgaagtgg acgtggactt acacatcggc cttcctggtt ttggtaaacc aagcaatgat 300
 gctaaacagc tgaagaagag aaatgggaag gagatcgcca catatgacgc cggaaaaggc 360
 60 atcgagaatg aactttccgg aaaggcatac tggatccccg cgccggagca aattctcata 420
 gggttcactc atttttcttg ccatgtatgc ttcaagacat tcaatcgcta caacaatctt 480

caggtaacgag tcaatatatc tcatgcgcat tgctttttcca tgcacaaaca tatataataa 540
attcatctta tagagttata tctccggatc taatgttatg agttttattca tatctatata 600
tatacatata tatatatata tatatatata tatatatata attctgaatt tatttgataa 660
5 aagctaaaca accaggattt aatagatgat ttaccttttg atctttattat acaatttaca 720
aatttaataca agtcaactaa tcgtgattta attacttttt tttgtaagaa gagttggtaa 780
tatatatattt tatggtaatg ttttcatgaa aataattcat cacaactctt tacattttatt 840
taatgcctta actaaagctg aattcgaaaa agttgaaata aattatctac taagatttga 900
ttgactatag tttttaatag ttttcttttc tcatatata attatcatag tagtcaaaac 960
10 atttgattca aacttaaata cacagatttc ttgaatgaaa cattactatg ctcggtcaat 1020
aatatgattt taaggaacca tgttatttca ttttattact taaggaaacc tttttgtttt 1080
ttgttgactc taaatattat gaatatagat gcacatgtgg ggacatgggt cacaatacag 1140
gaaaggaccg gagtcaactga aaggcacaca gccacgagcc atgttaggga tcccttgtaa 1200
ctgctgcgtt gaaggggtgca ggaaccacat tgaccatcct cgttccaagc cactgaaaga 1260
ctttaggacg ctccaaacgc actacaaacg caaacacgga cacaaaccct tctcgtgtcg 1320
15 cctttgcggt aagctttttg ctgtcaaggg cgattggcga acacatgaga agaattgtgg 1380
aaaacgttgg gtttgcgttt gcggttctga ttttaaacac aaacgttctc ttaaggacca 1440
tgtaaaggcg tttgggtctg gtcatgggccc ttatccaact ggtttgttt aagagcaggc 1500
ttctaattca tctgtctccg agactttgtt tttttaaatt tgggcatctt tttctttcgc 1560
ttatgaaata tctatttact ttagaaaaat aataatgtgg tatctaattg ttccaaatta 1620
20 ggaacacgaa gtgtaccatt atatttttca tcactacaaa tgttattcag agaaaattat 1680
cattaattgt ctcgttaaag atagaatagg gtttgaattt atcaaataat aaaaacagat 1740
caatacaaaa ttgaccatgc atatgcactt gaatattctg atttctttat gatgtaatct 1800
cattcaagaa aagctt 1816

25